

Why do tougher caseworkers increase employment? The role of programme assignment as a causal mechanism

Martin Huber*, Michael Lechner**, and Giovanni Mellace+

*University of Fribourg

**University of St. Gallen

+University of Southern Denmark

Abstract: Previous research found that less accommodating caseworkers are more successful in placing unemployed workers into employment. This paper explores the causal mechanisms behind this result using semiparametric mediation analysis. Analysing rich linked jobseeker-caseworker data for Switzerland, we find that the positive employment effects of less accommodating caseworkers are not driven by a particularly effective mix of labour market programmes, but rather by other dimensions of the counselling process, possibly including threat effects of sanctions and pressure to accept jobs.

Keywords: Unemployment, counselling style, active labour market policy, direct effects, indirect effects, causal mechanisms, causal channels, matching estimation

JEL classification: J64, J68, C21, C31.

We have benefitted from comments by the editor Bryan Graham, Judea Pearl, seminar/conference participants at the Universities of Melbourne, Sydney, and Otago, Copenhagen, Madrid (Carlos III), the EALE conference in Ljubljana, the workshop on Applied Microeconomics in Odense, the COMPIE conference in Rome, and the IAAE meeting in Thessaloniki, as well as three anonymous referees. Financial support from the Swiss National Science Foundation, grant number SNF 100018_137769 for the project “Identifying causal mechanisms in empirical economics” is gratefully acknowledged. **Addresses for correspondence:** Martin Huber (Martin.Huber@unifr.ch), Department of Economics, University of Fribourg, Bd. de Pérolles 90, CH-1700 Fribourg; Michael Lechner (Michael.Lechner@unisg.ch, www.michael-lechner.eu), Swiss Institute for Empirical Economic Research (SEW), University of St. Gallen, Varnbühlstrasse 14, CH-9000 St. Gallen; Giovanni Mellace (gione@sam.sdu.dk), Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M.

** Michael Lechner is also affiliated with CEPR and PSI, London, CESifo, Munich, IAB, Nuremberg, and IZA, Bonn.

1 Introduction

Caseworkers in employment offices usually have dual roles of jobseeker *counselling* and *monitoring* and may differ in performing these tasks. Behncke, Frölich, and Lechner (2010; henceforth BFL) found for Switzerland that less cooperative caseworkers increase the reemployment chances of unemployed. Reconsidering their linked jobseeker-caseworker data, this paper decomposes the average treatment effect on the treated (ATET) of having a less cooperative caseworker into an “indirect” effect through the assignment of an active labour market programme (ALMP) and a “direct” effect comprising all remaining causal mechanisms. That is, we consider the costly tool of ALMPs¹ as an explicit mediator of caseworkers’ counselling style. Applying a sequential conditional independence assumption and propensity score matching, our findings suggest that the total effect is driven by channels other than programme participation, which increase employment by initially roughly 1.5 percentage points, though the effect levels off over time. In contrast, the indirect path is never economically or statistically significant, such that the success of non-cooperative caseworkers does apparently not come from more effective ALMPs.

The literature on direct and indirect effects or mediation analysis (see Baron and Kenny, 1986) has recently moved towards flexible modelling, see for instance Flores and Flores-Lagunes (2009), Imai, Keele, and Yamamoto (2010), and Huber (2014). As methodological contribution, we propose a propensity score matching approach for estimating direct and indirect effects on the treated based on two matching steps using two different propensity scores: the conditional probabilities of treatment (i) given covariates and (ii) given covariates and mediators. On the empirical side, we are the first to evaluate the causal mechanisms of

¹In 2003, training programmes cost on average 110 CHF per person and day (plus additional expenses of 11 CHF for travelling etc.), employment programmes 98 CHF (plus 12 CHF), see Lalive, Zehnder, and Zweimüller, J. (2006).

caseworkers' placement success in a mediation framework using linked survey-administrative data on caseworkers and jobseekers.

The remainder is organized as follows. Section 2 discusses the econometric framework. Section 3 describes the data and selection issues. Section 4 presents the results.

2 Econometric framework

2.1 Potential outcomes and different causal effects

We aim at disentangling the effect of a binary treatment (D) on some outcome (Y) into a direct effect and an indirect effect operating through a possibly multidimensional mediator (M). We denote by $Y(d)$ and $M(d)$ the potential outcome and mediator states under treatment $d \in \{1, 0\}$ and by $\Delta_{D=1} = E[Y(1) - Y(0) | D=1]$ the ATET. To disentangle $\Delta_{D=1}$, rewrite the potential outcome as a function of both the treatment and the mediator: $Y(d) = Y(d, M(d))$. This allows formulating the direct effect ($\theta_{D=1}(d)$), i.e. the fraction of the total effect which is not attributed to the mediator, and the indirect effect ($\delta_{D=1}(d)$), i.e. the fraction attributed to the mediator, on the treated population:²

$$\begin{aligned}\theta_{D=1}(d) &= E[Y(1, M(d)) - Y(0, M(d)) | D=1], \\ \delta_{D=1}(d) &= E[Y(d, M(1)) - Y(d, M(0)) | D=1], d \in \{1, 0\}.\end{aligned}\tag{1}$$

Concerning the direct effect, Vansteelandt and VanderWeele (2012) argue that focussing on $M(1)$ appears to be the natural reference for treated subjects when the choice of reference levels is a priori hard to justify, because $M(1)$ corresponds to the actual choice of the treated:

² Robins and Greenland (1992) and Pearl (2001) use the denomination pure/total and natural direct and indirect effects, respectively.

$$\theta_{D=1}(1) = E[Y(1, M(1)) - Y(0, M(1)) | D=1] = E[Y - Y(0, M(1)) | D=1]. \quad (2)$$

As the ATET is the sum of the direct and indirect effects defined on opposite treatment states,

$\Delta_{D=1} = E[Y(1, M(1)) - Y(0, M(1)) | D=1] + E[Y(0, M(1)) - Y(0, M(0)) | D=1] = \theta_{D=1}(1) + \delta_{D=1}(0)$, the indirect effect $\delta_{D=1}(0)$ corresponds to the difference between the ATET and $\theta_{D=1}(1)$:

$$\delta_{D=1}(0) = \Delta_{D=1} - \theta_{D=1}(1) = E[Y(0, M(1)) - Y(0, M(0)) | D=1]. \quad (3)$$

In this paper, we therefore aim at identifying and estimating (2) and (3).³

2.2 Identifying assumptions and estimation

We impose sequential conditional independence of the treatment and the mediator:⁴

Assumption 1: $\{Y(0, m), M(0)\} \perp\!\!\!\perp x$ for all m and x in the common support.

Assumption 1 states that the joint distribution of the potential outcomes (for any m) and mediators under non-treatment are independent of the treatment conditional on X . This rules out unobserved confounders affecting the treatment on the one hand and the potential outcome and/or mediator under $D=0$ on the other hand, after controlling for observables.

Assumption 2: $Y(0, m) \perp\!\!\!\perp x, D=d$ for all m, d , and x in the common support.

By Assumption 2 the potential outcome under non-treatment and the observed mediator are independent conditional on the covariates and the treatment.⁵ This implies that $Y(0, m)$ is

³ The question asked by focussing on (2) and (3) is: “To which extent do the non-cooperative caseworkers’ current ALMP strategies and other mechanisms contribute to the overall effect?” This is different to: “How effective are specific combinations of caseworker types and programmes, given that policy makers may prescribe either?”, which may be addressed by multiple or dynamic treatment evaluation, see for instance Imbens (2000) and Lechner (2009). Answering the first question explores the causal mechanisms underlying the status quo, while answering the second may help improving upon the status quo.

⁴ Our identifying assumptions differ from those of Vansteelandt and VanderWeele (2012) who allow for confounding of M , but impose stronger assumptions on the potential outcomes.

⁵ Our assumptions require that $Y(0, M(1))$ and $Y(0, M(0))$ take the same value whenever $M(1) = M(0)$. This would be violated if the actual content of a formally identical ALMP differed across D conditional on X .

conditionally independent of (the common-world state) $M(0)$ whenever $D=0$ and of (the cross-world state) $M(1)$ whenever $D=1$.

Assumption 3: $\Pr(D=1 | M, X) < 1$.

By Assumption 3, there exists no combination of M, X that predicts treatment receipt with probability one.

Theorem 1: Under Assumptions 1 to 3, the following equalities hold:

$$E[Y(0, M(1)) | D=1] = E_{M, X | D=1} [E[Y | M, X, D=0]], \quad (4)$$

$$E[Y(0, M(0)) | D=1] = E_{X | D=1} [E[Y | X, D=0]], \quad (5)$$

with $E_{A|B=b}[C]$ denoting the expectation of C taken over the distribution of A conditional on $B=b$.

Thus, $\theta_{D=1}(1)$ and $\delta_{D=1}(0)$ are identified. Proof: See Appendix A.

Our Theorem 1 and its proof are similar to Theorem 1 of Imai, Keele, and Yamamoto (2010), albeit the latter focus on the total population, while we consider the treated. Instead of conditioning on M and X as in (4) and (5), the results of Rosenbaum and Rubin (1983) imply that identification is also obtained conditional on the propensity scores $p_{mx}(M, X) = \Pr(D=1 | M, X)$ and $p_x(X) = \Pr(D=1 | X)$, which balance the distributions of (M, X) and X , respectively:⁶

$$\begin{aligned} \theta_{D=1}(1) &= E[Y | D=1] - E_{p_{mx}(M, X) | D=1} [E[Y | p_{mx}(M, X), D=0]], \\ \delta_{D=1}(0) &= E_{p_{mx}(M, X) | D=1} [E[Y | p_{mx}(M, X), D=0]] - E_{p_x(X) | D=1} [E[Y | p_x(X), D=0]]. \end{aligned}$$

We therefore propose estimating the direct and indirect effects by propensity score-based matching of non-treated observations to the treated (i) using the estimated $p_{mx}(M, X)$ to

⁶ Specifically, $X \perp\!\!\!\perp D | (M, X)$ and $(M, X) \perp\!\!\!\perp D | p_{mx}(M, X)$ for all m, x in the common support.

obtain an estimate of $\theta_{D=1}(1)$ and (ii) using the estimated $p_x(X)$ to estimate the ATET from which the direct effect is subtracted to estimate $\delta_{D=1}(0)$.⁷

We use distance-weighted radius matching on the propensity score with bias adjustment (see Lechner, Miquel, and Wunsch, 2011) and model $p_x(X)$ and $p_{mx}(M, X)$ by flexible probit specifications. Concerning inference, we apply block bootstrapping and (i) resample caseworkers along with their jobseekers to account for clustering, (ii) compute the bootstrap t-statistics of the effects, and (iii) estimate the p-value as the smoothed share of absolute bootstrap t-statistics that exceeds the absolute t-statistic in the original sample, see Racine and MacKinnon (2007).

3 Empirical implementation

3.1 Study sample and definition of treatment, mediators, and outcomes

The study sample comprises individuals who registered at Swiss employment offices in 2003. Our administrative data include among others nationality, qualification, education, experience, profession and industry of last job, participation in ALMPs, employment histories, and regional characteristics. They were linked to a caseworker survey on aims, strategies, and processes of employment offices and caseworkers. The sample definition follows BFL (see Appendix B.1.) and ultimately contains 1,284 caseworkers and 100,120 unemployed.

The survey contained a question on the importance of cooperation with the jobseeker, see Table 2.1.⁸ As in BFL, the treatment (D) is one if the caseworker chose option 2 or 3 (is

⁷ Without interactions between D and M in the outcome model, the indirect effect could alternatively be identified as the effect of D on M multiplied with the effect of M on Y . This approach could incorporate post-treatment confounders influenced by D that affect both M and Y . There exists thus a trade-off between model flexibility and allowing for post-treatment confounding, see for instance Imai and Yamamoto (2013).

less cooperative) and zero otherwise.⁹ The mediator (M) is the *first* participation in an ALMP within six months after starting the unemployment spell. We categorize ALMPs into six mutually exclusive groups: job search training, personality course, language skill training, computer training, vocational training, and participation in an employment programme/internship. Together with non-participation, this entails seven possible mediator states. Estimation is based on six dummies for each ALMP being the first programme in 2003. Table B2.1 in Appendix B.2 provides the frequencies of ALMPs across treatment states.

Table 2.1: Survey question on cooperativeness of the caseworker

How important do you consider the cooperation with the jobseeker, regarding placements and assignment of active labour market programmes?	
<input type="checkbox"/> 1	Cooperation is very important; the wishes of the unemployed person should be satisfied.
<input type="checkbox"/> 2	Cooperation is important, but placements and ALMP should sometimes be assigned or declined in spite of unemployed person's wishes.
<input type="checkbox"/> 3	Cooperation is less important; I should assign placements and ALMP independent of the wishes of the unemployed person
Note: 52% of the caseworkers chose option one, 39% of caseworkers chose option two, and 9% of caseworkers chose option three. Only very few caseworkers did not respond to this particular question. They are dropped from the analysis.	

The outcome (Y) is a binary employment indicator on monthly bases. An individual is considered employed if she has de-registered at the employment office and the exit state is employment. As the mediator causally precedes the outcome, we only consider employment states assessed from month 8 after caseworker assignment (at least one month after programme start). At least 29 outcome periods are available: At the latest, jobseekers are assigned to a caseworker in the end of 2003 so that the mediators are measured from the beginning until the first half of 2004 and the outcome from the second half of 2004 until the

⁸ According to the survey, non-cooperative caseworkers consider control and sanctions, job assignments, and employment programmes as more important, and counselling meetings and temporary wage subsidies as less important. They also stated to assign ALMPs to exert pressure and to control their clients' availability for jobs.

⁹ The self-reported treatment may contain measurement error with respect to the actual cooperativeness. For instance, some caseworkers might state to be (non-)cooperative but act differently in practice and this may be correlated with unobserved characteristics like personal integrity that also affect the outcomes. We nevertheless believe that measurement error is limited, as there are no obvious incentives for misreporting (like justification bias).

end of in 2006. We therefore estimate the effects over a period of 1.5 to 3 years after caseworker assignment.

3.2 Selectivity of the treatment and the mediators

Our strategy requires observing all confounders of the relationship of the outcome with both the treatment and/or the mediator (conditional on treatment). BFL argue that treatment selection depends on the types of hired caseworkers, the assignment of unemployed to caseworkers, and the development of the caseworkers' attitudes as they gain work experience. As attitudes may be related to job placement success, we control for caseworker characteristics like age, gender, education, work experience, and own unemployment experience. We also control for rules assigning the unemployed to caseworkers, as stated in the survey. Furthermore, caseworkers may react differently to various types of unemployed and the labour market environment. Therefore, we include a range of jobseeker characteristics like gender, qualification, previous occupation, labour market history, and a caseworker-provided employability rating, as well as local labour market conditions.

Nevertheless, some individual characteristics like skills not reflected in qualifications, attitude, and motivation are unobserved, but arguably associated with observables. For instance, motivation should correlate with the labour market history and the employability rating (which should furthermore also capture unobserved skills). We therefore believe that conditional on observables, variation in the assignment of caseworker types is likely exogenous and for instance related to the current availability and work load of caseworkers.

Characteristics that have been identified as controls for caseworker cooperativeness are also expected to influence selection into the programmes. As argued by Gerfin and Lechner (2002) and many others, the assignment of ALMPs is likely driven by the jobseeker's socio-economic background, previous occupation and labour market experience, employability, and

regional characteristics. In addition, caseworker characteristics might affect the assignment strategy. Conditional on these factors, exogenous variation in ALMPs likely comes from differences in their availability across employment offices. Finally, it is worth noting that programme assignment usually takes place early in the unemployment spell, so that time-varying confounding of the mediators due to changes in relevant factors during the unemployment spell is likely a minor issue. The fact that we use a (short) six months window for programme start is also in favour of this argument.

Table B.2.2 in Appendix B.2 provides descriptive statistics for a range of confounders by treatment and programme status. While there is limited selection with respect to caseworker rigour, perhaps with the exception of regional aspects, selection into the programmes is much stronger and driven by several factors. See also Table B.3 of Appendix B.3 for the propensity score specifications. Appendix C contains tests for whether the propensity scores balance the characteristics of treated and non-treated in matching estimation (Tables C.1 and C.2) and suggest that balancing works well.

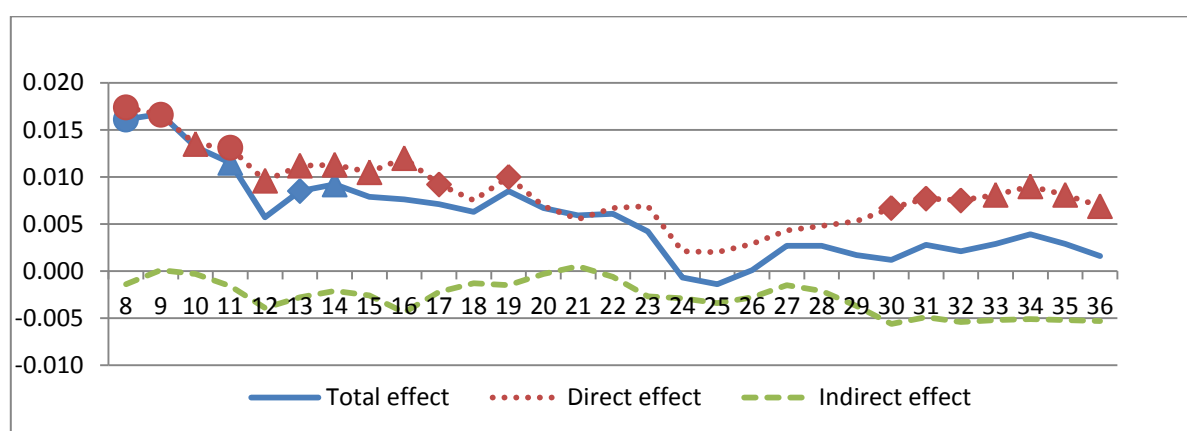
4 Results

Figure 6.1 contains the ATET as well as the direct and indirect effects among the treated on employment from month 8 to 36. The three lines represent the total, direct, and indirect effects and superimposed symbols imply that these particular effects are (pointwise) significant at the indicated level.¹⁰ The results suggest that initially, less cooperative caseworkers significantly increase employment by roughly 1.5%-points. Over time, however, the ATET vanishes and is not statistically significant after month 14, even though it remains positive in almost all months. The (initial) employment gain is mainly driven by the direct effect of caseworker rigour, the indirect mechanism through ALMPs is never significant.

¹⁰ Appendix D provides the results for the outcomes unemployment with benefit receipt and looking for a job.

Our estimates therefore suggest that the (at least initially) higher job placement rates of less cooperative caseworkers are not due to better programmes, but rather to other dimensions that possibly include the threat of sanctions¹¹ or pressure to go to job interviews. This suggests that policy makers should not only be interested in the effective provision of ALMPs, but also in the analysis of other dimensions of caseworkers' counselling style, which can apparently make a difference.

Figure 6.1: Effects on employment by month after registration



Note: A route / triangle / circle implies pointwise significance at the 10% / 5% / 1% level, respectively.

References

- Baron, R.M., and Kenny, D.A. (1986): The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Behncke, S., Frölich, M., and Lechner, M. (2010): Unemployed and their caseworkers: should they be friends or foes? *Journal of the Royal Statistical Society Series A*, 173, 67–92.
- Flores, C.A., and Flores-Lagunes, A. (2009): Identification and Estimation of Causal Mechanisms and Net Effects of a Treatment under Unconfoundedness. IZA Discussion Paper No. 4237.
- Gerfin, M., and Lechner, M. (2002): Microeconomic Evaluation of the Active Labour Market Policy in Switzerland. *Economic Journal*, 112, 854–893.
- Huber, M. (2014): Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, 29, 920–943.
- Imai, K., Keele, L., and Yamamoto, T. (2010): Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, 25, 51–71.

¹¹ We observe actual sanctioning days but do not find any economically or statistically significant effect of counselling style on sanctioning. Cooperative and non-cooperative caseworkers could nevertheless differ in terms of the (unobserved) threat of sanctions, which may also affect employment.

- Imai, K., and Yamamoto, T. (2013): Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments. *Political Analysis*, 21, 141–171.
- Imbens, G. W. (2000): The Role of the Propensity Score in Estimating Dose-response Functions. *Biometrika*, 87, 706–710.
- Lalive, R., Zehnder, T., and Zweimüller, J. (2006): Makroökonomische Evaluation der Aktiven Arbeitsmarktpolitik der Schweiz. Swiss Federal Department of Economics - Publication in Labour Market Policy No 19 (Seco Publikation Arbeitsmarktpolitik No 19).
- Lechner, M. (2009): Sequential Causal Models for the Evaluation of Labor Market Program. *Journal of Business & Economic Statistics*, 27, 71–83.
- Lechner, M., Miquel, R., and Wunsch, C. (2011): Long-Run Effects Of Public Sector Sponsored Training In West Germany. *Journal of the European Economic Association*, 9, 742–784.
- MacKinnon, J.G. (2006): Bootstrap Methods in Econometrics. *The Economic Record*, 82, S2–S18.
- Pearl, J. (2001): Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, San Francisco: Morgan Kaufman, 411–420.
- Racine, J. S., and MacKinnon, J. G. (2007): Inference via kernel smoothing of bootstrap values. *Computational Statistics & Data Analysis*, 51, 5949–5957.
- Robins, J.M., and Greenland, S. (1992): Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*, 3, 143–155.
- Rosenbaum, P., and Rubin, D.B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41–55.
- Vansteelandt, S. and VanderWeele, T.J. (2012). Natural direct and indirect effects on the exposed: effect decomposition under weaker assumptions. *Biometrics*, 68, 1019–1027.